MA-GY 7043: Linear Algebra II

Complex Tensors
Symmetric and Hermitian Forms
Cayley-Hamilton Theorem
Spectral Mapping Theorem
Nilpotent Matrices
Generalized Eigenspaces

Deane Yang

Courant Institute of Mathematical Sciences
New York University

April 29, 2025

Outline I

$(1,0)$-Tensors and $(0,1)$-Tensors on Complex Vector Space

- Let V be a complex vector space
- A $(1,0)$-tensor is a linear function $\ell : V \to \mathbb{C}$, i.e.,

$$\ell(av + bw) = a\ell(v) + b\ell(w)$$

- Let $V^{(1,0)}$ denote the space of all $(1,0)$-tensors on V
- A $(0,1)$-tensor is a conjugate linear function $\ell : V \to \mathbb{C}$, i.e.,

$$\ell(av + bw) = \bar{a}\ell(v) + \bar{b}\ell(w)$$

- Let $V^{(0,1)}$ denote the space of all $(0,1)$-tensors on V
- Both $V^{(1,0)}$ and $V^{(0,1)}$ are complex vector spaces
- If ℓ is a $(1,0)$-tensor, then $\bar{\ell}$ is a $(0,1)$-tensor and therefore the map

$$V^{(1,0)} \to V^{(0,1)}$$
$$\ell \mapsto \bar{\ell}$$

is a conjugate linear isomorphism

(k, l)-tensors on Complex Vector Space

- Given nonnegative integers k, l such that $k + l \leq \dim(V)$, a function

$$V \times \cdots \times V \to \mathbb{C}$$

$$(v_1, \ldots, v_k, v_{k+1}, \cdots, v_{k+l}) rightarrow T(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l})$$

is a (k, l)-tensor if for each $1 \leq i \leq k$, the function

$$v_i \mapsto T(v_1, \ldots, v_{i-1}, v_i, \ldots, v_k, \ldots, v_{k+l})$$

is linear and for each $k + 1 \leq i \leq l$, the function

$$v_i \mapsto T(v_1, \ldots, v_k, \ldots, v_{i-1}, v_i, \ldots, v_{k+l})$$

is conjugate linear

Sesquilinear and Quadratic Forms on Complex Vector Space

- A **sesquilinear form** on a complex vector space V is an element of $V^{1,1}$
- I.e., a function
$$B : V \times V \to \mathbb{C}$$
with the following properties: For all $v, v_1, v_2 \in V$,
$$\begin{aligned} B(v_1 + v_2, v) &= B(v_1, v) + B(v_2, v) \\ B(v, v_1 + v_2) &= B(v, v_1) + B(v, v_2) \\ B(cv_1, v_2) &= cB(v_1, v_2) \\ B(v_1, cv_2) &= \bar{c}B(v_1, v_2) \end{aligned}$$
- A **quadratic form** is a fumction $Q : V \to \mathbb{C}$ for which there exists a sesquilinear form B such that for any $v \in V$,
$$Q(v) = B(v, v)$$

Hermitian Forms

- A sesquilinear form H is **hermitian** if for any $v_1, v_2 \in V$,

$$H(v_2, v_1) = \overline{H(v_1, v_2)}$$

- In particular, for any $v \in V$, setting $v_1 = v_2 = v$ gives

$$H(v, v) = \overline{H(v, v)},$$

  which implies $H(v, v) \in \mathbb{R}$

- Therefore, the quadratic form associated with a hermitian form H is a real-valued function on V,

$$Q : V \to \mathbb{R}$$

Tensor Product of $(1,0)$-Tensor with $(0,1)$-Tensor

- Given $\theta^1 \in V^{(1,0)}$ and $\theta^2 \in V^{(0,1)}$, define the function

$$\theta^1 \otimes \theta^2 : V \times V \to \mathbb{C}$$

given by

$$(\theta^1 \otimes \theta^2)(v_1, v_2) = \langle \theta^1, v^1 \rangle \langle \theta^2, v^2 \rangle$$

- Observe that $\theta^1 \otimes \theta^2$ is sesquilinear because θ^1 is linear and θ^2 is conjugate linear
- The $(1,1)$-tensor

$$\theta^1 \otimes \theta^2 + \bar{\theta}^2 \otimes \bar{\theta}^1$$

is hermitian

Sesquilinear Forms With Respect to Basis

- Let $(e_1, \ldots, e_m)$ be a basis of V and $(\epsilon^1, \ldots, \epsilon^m) \subset V^{(1,0)}$ be the dual basis
- Let $B \in V^{(1,1)}$ and for each $1 \leq i, j \leq m$,

$$M_{ij} = B(e_i, e_j)$$

- For any $v = e_i a^i$ and $w = e_j b^j$,

$$\begin{aligned} B(v, w) &= B(e_i a^i, e_j b^j) \\ &= a^i \bar{b}^j B(e_i, e_j) \\ &= a^i \bar{b}^j M_{ij} \end{aligned}$$

- Conversely, any matrix $M \in \mathrm{gl}(m, \mathbb{C})$ defines a sesquilinear form using the same formula
- This defines a bijective linear map $V^{(1,1)} \to \mathrm{gl}(m, \mathbb{C})$

Basis of $V^{(1,1)}$

- On the other hand,

$$\begin{aligned}(\epsilon^i \otimes \bar{\epsilon}^j)(v, w) &= (\epsilon^i \otimes \bar{\epsilon}^j)(e_k a^k, e_l b^l) \\ &= a^k b^l \langle \epsilon^i, e_k \rangle \langle \epsilon^j, e_l \rangle \\ &= a^k b^l \delta_{kl} \\ &= a^i b^j\end{aligned}$$

- Therefore, for any $v, w \in V$,

$$B(v, w) = M_{ij}(\epsilon^i \otimes \epsilon^j)(v, w),$$

i.e.,

$$B = M_{ij}(\epsilon^i \otimes \epsilon^j)$$

- Moreover, $B = 0 \iff \forall 1 \leq i, j \leq m, \; M_{ij} = 0$
- It follows that

$$\{\epsilon^i \otimes \epsilon^j : \; 1 \leq i, j \leq m\}$$

is a basis of $V^{(1,1)}$

Change of Basis Formula for Hermitian Form

- $H \in V^{(1,1)}$ be hermitian
- Given a basis $(e_1, \ldots, e_m)$ of V, let

$$M_{ij} = H(e_i, e_j)$$

- Given a basis $(f_1, \ldots, f_m)$ of V, let

$$N_{ij} = H(f_i, f_j)$$

- If $f_i = e_j A^j_i$, then

$$\begin{aligned} N_{ij} &= H(f_i, f_j) \\ &= H(e_k A^k_i, e_l A^l_j) \\ &= A^k_i \bar{A}^l_j H(e_k, e_l) \\ &= A^k_i M_{kl} \bar{A}^l_j, \end{aligned}$$

i.e.,

$$N = AM\bar{A}^T = AMA^*$$

Normal Form of Hermitian Form (Part 1)

- Recall that if a matrix M is hermitian, there exists a unitary matrix U such that $D = U^* M U$ is diagonal and the diagonal entries (eigenvalues of M) are real
- Let E be the diagonal matrix whose diagonal entries are

$$E_{kk} = \begin{cases} |D_{kk}|^{-1/2} & \text{if } D_{kk} \neq 0 \\ 1 & \text{if } D_{kk} = 0 \end{cases}$$

- Observe that $E^* D E = EDE$ is a diagonal matrix where

$$(E^* D E)_{kk} = \begin{cases} 1 & \text{if } D_{kk} > 0 \\ -1 & \text{if } D_{kk} < 0 \\ 0 & \text{if } D_{kk} = 0 \end{cases}$$

- If $V = UE$ and $N = V^* M V$, then

$$N = V^* M V = E^* U^* M U E = E^* D E$$

Normal Form of Hermitian Form (Part 2)

- With respect to the basis $(f_1, \dots, f_m)$ where

$$f_j = e_i N^i_j,$$

we get

$$H(f_i, f_j) = \begin{cases} \delta_{ij} & \text{if } D_{ii} > 0 \\ -\delta_{ij} & \text{if } D_{ii} < 0 \\ 0 & \text{if } D_{ii} = 0 \end{cases}$$

Signature of Hermitian Form

- The **signature** of a diagonal matrix is (a, b, c), where a is the number of positive diagonal elements, b is the number of negative diagonal elements, and c is the number of zero diagonal elements
- The **signature** of a hermitian matrix is (a, b, c), where a is the number of positive eigenvalues, b is the number of negative eigenvalues, and c is the number of zero eigenvalues

Sylvester's Law of Inertia

- Let $H \in V^{(1,1)}$ be a hermitian form on a complex vector space V
- Let $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ be bases of V that both diagonalize H
- Let M be the diagonal matrix given by

$$M_{ij} = H(e_i, e_j)$$

- Let N be the hermitian matrix given by

$$N_{ij} = H(f_i, f_j)$$

- **Theorem.** M and N have the same signature
- We can therefore define the **signature** of a hermitian form to be the signature of the hermitian matrix associated with a basis of V

Proof (Part 1)

- We can assume that M and N are diagonal where each diagonal element is 1, -1, or 0
- Let r be the number of positive values in $\{M_{11}, \ldots, M_{mm}\}$
- By permuting the basis vectors $e_1, \ldots, e_m$, we can assume that

$$M_{11} = H(e_1, e_1), \ldots, M_{rr} = H(e_r, e_r)$$

  are all positive
- Let R be the subspace spanned by $(e_1, \ldots, e_r)$
- Let s be the number of positive values in $\{N_{11}, \ldots, N_{mm}\}$
- By permuting the basis vectors $f_1, \ldots, f_m$, we can assume that

$$N_{11} = H(f_1, f_1), \ldots, N_{rr} = H(f_s, f_s)$$

  are all positive
- Let S be the subspace spanned by $\{f_1, \ldots, f_s\}$

Proof (Part 2)

- Define the projection map

$$P : V \to R$$
$$v = e_1 v^1 + \cdots + e_n v^n \mapsto e_1 v^1 + \cdots + e_r v^r$$

- Let $P_S : S \to R$ be the restriction of P to S
- Let $Q : V \to \mathbb{R}$ be the quadratic form where

$$Q(v) = H(v, v)$$

- On one hand, if $v \in S$, then $v = f_1 b^1 + \cdots + f_s b^s$ and

$$Q(v) = Q(f_1 b^1 + \cdots + f_s b^2) = \beta_1 (b^1)^1 + \cdots + \beta_s (b^s)^2 > 0$$

- On the other hand, if $v \in \ker P_S$, then

$$v = e_{r+1} a^{r+1} + \cdots + e_n a^n$$

and therefore

$$Q(v) = \alpha_{r+1} (a^{r+1})^2 + \cdots + \alpha_n (a^n)^2 \leq 0$$

Proof (Part 3)

- It follows that $\ker(P_S) = \{0\}$ and therefore $s = \dim(S) \leq r = \dim(R)$
- The same argument with the bases switched implies that $r = \dim(R) \leq s = \dim(S)$
- The same argument proves that the number of negative values in $\{M_{11}, \ldots, M_{mm}\}$ is equal to the number of negative values in $\{N_{11}, \ldots, N_{mm}\}$
- It follows that the signatures of M and N are equal

Bilinear and Sesquilinear Forms as a Linear Maps

- If B is a bilinear form on a real vector space V, then it defines a linear map
$$L_B : V \to V^*,$$
where for each $v, w \in V$,
$$\langle L_B(v), w \rangle = B(v, w)$$
- If B is a sesquilinear form on a complex vector space V, then it defines a linear map
$$L_B : V \to V^{(0,1)},$$
where for each $v, w \in V$,
$$\langle L_B(v), w \rangle = B(v, w)$$
and a conjugate linear map
$$R_B : V \to V^{(1,0)},$$
where for each $v, w \in V$,
$$\langle w, R_B(v) \rangle = B(w, v)$$

Linear Map of Bilinear or Sesquilinear Form With Respect to Basis

- Let $(e_1, \dots, e_m)$ be a basis of V and $(\epsilon^1, \dots, \epsilon^m)$ be the dual basis
- If $M_{ij} = B(e_i, e_j)$, then

$$\langle e_j, L_B(e_i)\rangle = B(e_i, e_j) = M_{ij}$$

- Therefore, $L_B(e_i) = M_{ij}\epsilon^j$
- It follows that if $v = e_i a^i \in \ker(L_B)$, then

$$0 = L_B(v) = L_B(e_i a^i) = L_B(e_i)a^i = (M_{ij}a^i)\epsilon^j,$$

i.e.,

$$Ma = 0$$

Degenerate and Nondegenerate Bilinear Forms

- A bilinear or sesquilinear form $B : V \times V \to \mathbb{R}$ is **degenerate** if there exists $v \neq 0$ such that for any $w \in V$,

$$B(v, w) = 0,$$

i.e.,

$$L_B(v) = 0$$

- A bilinear form B on a real vector space V is **nondegenerate** if it is not degenerate, i.e.,

$$\ker(L_B) = \{0\},$$

or, equivalently,

$$L_B : V \to V^*$$

is an isomorphism

- It follows that B is nondegenerate if and only if M is an invertible matrix

Signature of Nondegenerate Symmetric or Hermitian Form

- Recall that if an m-by-m symmetric or hermitian matrix M has signature (p, q), then

$$\dim(\ker(M)) = m - p - q$$

- It follows that M is invertible if and only if $p + q = m$
- It follows that a symmetric or hermitian form H is nondegenerate if and only if the its signature (p, q) satisfies

$$p + q = m$$

Different Notation Conventions for Hermitian Form

- We are using the following convention:

$$B(cv, w) = cB(v, w)$$
$$B(v, cw) = \bar{c}B(v, w)$$

- Some use the following convention:

$$B(cv, w) = \bar{c}B(v, w)$$
$$B(v, cw) = cB(v, w)$$

- When reading a paper or book, look carefully to see which convention is used

Cayley-Hamilton Theorem

- Recall that the characteristic polynomial of a square matrix A is

$$p(x) = \det(A - xI)$$

- Given any polynomial

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

and square matrix M, we can define

$$p(M) = a_0 I + a_1 M + \cdots + a_n M^n$$

- **Theorem:** If p is the characteristic polynomial of a square matrix A, then

$$p(M) = 0$$

Wrong Proof

- Since $p(x) = \det(A - xI)$,

$$p(A) = \det(A - AI) = 0$$

Characteristic Polynomial

- Recall that if A is a square polynomial over $\mathbb{C}$, its characteristic polynomial is

$$p_A(x) = \det(A - xI) = (\lambda_1 - x) \cdots (\lambda_n - x),$$

  where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, counting multiplicities
- Therefore, for each eigenvalue λ_k,

$$p_A(\lambda_k) = 0$$

Polynomial Function of Diagonal Matrix (Part 1)

- Given a polynomial

  $$p(x) = a_0 + a_1 x + \cdots + a_k x^k,$$

  and a diagonal matrix,

  $$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

  let

  $$p(D) = a_0 I + a_1 D + \cdots + a_n D^n$$

Polynomial Function of Diagonal Matrix (Part 2)

- Therefore,

$$
\begin{aligned}
& p(D) \\
& = a_0 I + a_1 \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} + \cdots + a_n \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{bmatrix} \\
& = \begin{bmatrix} a_0 + a_1\lambda_1 + \cdots + a_n\lambda_1^n & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & a_0 + a_1\lambda_n + \cdots + a_n\lambda_n^n \end{bmatrix} \\
& = \begin{bmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & p(\lambda_n) \end{bmatrix}
\end{aligned}
$$

Proof of Cayley-Hamilton For Diagonal Matrix

- Therefore,

$$p_D(D) = \begin{bmatrix} p_D(\lambda_1) & 0 & \cdots & 0 \\ 0 & p_D(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & p_D(\lambda_n) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Cayley-Hamilton For Diagonalizable Matrix (Part 1)

- If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then since
$$0 = p_A(\lambda_k) = \det(A - \lambda_k I)$$
- If A is diagonalizable, then there is an invertible matrix M such that
$$A = MDM^{-1},$$
where
$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

Cayley-Hamilton For Diagonalizable Matrix (Part 2)

- Observe that for each positive integer k,

$$
\begin{aligned}
(MDM^{-1})^k &= (MDM^{-1}) \cdots (MDM^{-1}) \\
&= MD(M^{-1}M) \cdots D(M^{-1}M)DM^{-1} \\
&= MD^k M^{-1}
\end{aligned}
$$

- Observe that

$$
\begin{aligned}
p_A(x) &= \det(A - xI) \\
&= \det(MDM^{-1} - M(xI)M^{-1}) \\
&= (\det(M)) \det(D - xI)(\det(M^{-1}) \\
&= \det(D - xI) = p_D(x)
\end{aligned}
$$

- Therefore,

$$
\begin{aligned}
p_A(A) &= a_0 I + a_1 A + \cdots + a_n A^n \\
&= a_0 MIM^{-1} + a_1 MDM^{-1} + \cdots + a_n (MDM^{-1})^n \\
&= M(a_0 I + a_1 D + \cdots + a_n D^n)M^{-1} \\
&= Mp_D(D)M^{-1}
\end{aligned}
$$

Proof of Cayley-Hamilton Using Analysis

- For any square matrix A, there exists a sequence of diagonalizable matrices that converges to A
- The map

$$\begin{aligned} \mathrm{gl}(n, \mathbb{F}) \times \mathrm{gl}(n, \mathbb{F}) &\to \mathrm{gl}(n, \mathbb{F}) \\ (A, B) &\mapsto p_A(B) \end{aligned}$$

  is continuous
- Therefore,

$$p_A(A) = \lim_{k \to \infty} p_{A_k}(A_k) = 0$$

Abstract Cayley-Hamilton Formula

- Recall the characteristic polynomial of a linear map $A : V \to V$ is given by

$$p_A(x) = \det(A - xI) = (-1)^n(x - \lambda_1)\cdots(x - \lambda_n) = a_0 + a_1x + \cdots + a_nx^n$$

  where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, counting multiplicities
- Then

$$p_A(A) = (-1)^n(A - \lambda_1)\cdots(A - \lambda_n) = a_0I + a_1A + \cdots + a_nA^n = 0$$

- Since $p_A(A)$ is a linear map from V to V, this is equivalent to saying that for any $v \in V$,

$$p_A(A)v = 0$$

Proof Using Schur Decomposition (Part 1)

- Let $A : V \to V$ have eigenvalues $\lambda_1, \ldots, \lambda_n$, counting multiplicities
- Then there exists a basis $(e_1, \ldots, e_n)$ of V such that for each $1 \le k \le n$,
  $$A(e_k) = M_k^1 e_1 + \cdots + M_k^k e_k,$$
  where $M_k^k = \lambda_k$
- Let E_k be the span of $\{e_1, \ldots, e_k\}$
- Observe that $A(E_k) \subset E_k$
- Since
  $$\begin{aligned}(A - \lambda_k I)e_k &= M_k^1 e_1 + \cdots + M_k^{k-1} e_{k-1} + (M_k^k - \lambda_k)e_k \\ &= M_k^1 e_1 + \cdots + M_k^{k-1} e_{k-1} \\ &\in E_{k-1},\end{aligned}$$
  it follows that
  $$(A - \lambda_k I)(E_k) \subset E_{k-1}$$

Proof Using Schur Decomposition (Part 2)

- Therefore, for any $v \in V = E_n$,

$$\begin{aligned} (A - \lambda_n I)v &\in E_{n-1} \\ (A - \lambda_{n-1} I)(A - \lambda_n I)v &\in E_{n-2} \\ \vdots \quad & \quad \vdots \\ (A - \lambda_2 I) \cdots (A - \lambda_n I)v &\in E_1 \\ (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I)v &= 0 \end{aligned}$$

- Therefore, for any $v \in V$,

$$p_A(A)v = (A - \lambda_1 I) \cdots (A - \lambda_n I)v = 0$$

- It follows that $p_A(A) = 0$

Spectral Mapping Theorem

- Given a polynomial p and a subset $S \subset \mathbb{C}$, let

$$p(S) = \{p(z) \ : \ z \in S\}$$

- Let V be a complex vector space and $L : V \to V$ be a linear map
- The spectrum of L, denoted $\sigma(L)$, is the set of all eigenvalues of L, not counting multiplicity
- **Theorem.** For each linear map $L : V \to V$ and polynomial p,

$$\sigma(p(L)) = p(\sigma(L))$$

- **Corollary.** $p(L)$ is invertible if and only if $0 \notin p(\sigma(L))$

$p(\sigma(L)) \subset \sigma(p(L))$

- If $v \in V$ is an eigenvector of L with eigenvalue λ, then

$$\begin{aligned} Lv &= \lambda v \\ L^2 v &= L(Lv) = L(\lambda v) = \lambda^2 v \\ L^k v &= \lambda^k v \end{aligned}$$

- Therefore, if $p(x) = a_0 + a_1 x + \cdots + a_k x^k$, then

$$\begin{aligned} p(L)v &= (a_0 + a_1 L + \cdots + a_k L^k)v \\ &= a_0 v + a_1 Lv + \cdots + a_k L^k v \\ &= a_0 + a_1 \lambda v + \cdots + a_k \lambda^k v \\ &= p(\lambda)v \end{aligned}$$

- It follows that for each eigenvalue λ of L,

$$p(\lambda) \in \sigma(p(L))$$

and therefore

$$p(\sigma(L)) \subset \sigma(p(L))$$

$\sigma(p(L)) \subset p(\sigma(L))$ (Part 1)

- Let $\mu \in \sigma(p(L))$ and v be a corresponding eigenvector
- Let $q(z) = p(z) - \mu$, which implies

$$q(L) = p(L) - \mu I$$

- Then $q(L) : V \to V$ is not invertible, because

$$q(L)v = p(L)v - \mu v = 0$$

- By the Fundamental Theorem of Algebra, q can be factored

$$q(z) = a_k(z - z_1)\cdots(z - z_k),$$

  where $z_1, \ldots, z_k$ are the roots of q, counted with multiplicity
- Therefore, $q(L) = a_k(L - z_1)\cdots(L - z_k)$

$\sigma(p(L)) \subset p(\sigma(L))$ (Part 2)

- Since
$$q(L) = a_k(L - z_1 I) \cdots (L - z_k I)$$
is not invertible, at least one of the factors $L - z_j I$ is not invertible
- It follows that $z_j \in \sigma(L)$
- Since
$$p(z_j) = q(z_j) + \mu = \mu \in \sigma(p(L)),$$
it follows that for each $\mu \in \sigma(p(L))$, there exists $\lambda \in \sigma(L)$ such that
$$p(\lambda) = \mu$$
- Therefore, $\sigma(p(L)) \subset p(\sigma(L))$

Example of Nilpotent Matrix

- Consider the following example:

$$M_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

- Since

$$\det(M_0 - \lambda I) = \det\left(\begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix}\right) = \lambda^2,$$

  the only eigenvalue of M_0 is 0
- On the other hand, v is an eigenvector if and only if

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = M_0 v = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = \begin{bmatrix} v^2 \\ 0 \end{bmatrix}$$

- Therefore, the eigenspace for $\lambda = 1$ is only 1-dimensional
- On the other hand,

$$M_0^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Another Nilpotent Matrix

- Consider the following example:

$$M_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$M_0^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$M_0^3 = M_0^2 M_0 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Example

- Consider the following example:

$$M_\lambda = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$M_\lambda - \lambda I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$(M_\lambda - \lambda I)^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$(M_\lambda - \lambda I)^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Diagonalizable Example

- On the other hand, if $\lambda_1, \lambda_2, \lambda_3$ are distinct, then

$$M = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

is diagonalizable

Abstract Description of Nilpotent Matrix

- If (e_1, e_2, e_3) is the standard basis of $\mathbb{R}^3$, then

$$M_0 e_1 = e_2$$
$$M_0 e_2 = e_3$$
$$M_0 e_3 = 0$$

and

$$(M_\lambda - \lambda I)e_1 = e_2$$
$$(M_\lambda - \lambda I)e_2 = e_3$$
$$M_\lambda - \lambda I)e_3 = 0$$

Generalized Eigenspaces (Part 1)

- Consider a linear map $L : V \to V$
- $v \in V$ is a **generalized eigenvector** of L for the eigenvalue λ if there exists $k \in \mathbb{Z}^+$ such that

$$(L - \lambda I)^k v = 0$$

- The generalized eigenspace of λ is the set E_λ of all generalized eigenvectors along with 0,

$$E_\lambda = \bigcup_{k \geq 1} \ker((L - \lambda I)^k)$$

- This is a nested sequence of subspaces

$$\ker(L - \lambda I) \subset \ker((L - \lambda I)^2) \subset \cdots$$

Generalized Eigenspaces (Part 2)

- The sequence cannot be infinite and therefore there exists k such that
$$\ker((L-\lambda I)^k) = \ker((L-\lambda I)^{k+1})$$
- Therefore, if $v \in \ker(L-\lambda I)^{k+l+1}$, then
$$(L-\lambda I)^l v \in \ker(L-\lambda I)^{k+1} = \ker(L-\lambda I)^k,$$
which implies
$$(L-\lambda I)^{k+l} v = (L-\lambda I)^k (L-\lambda I)^l v = 0$$
- So
$$v \in \ker(L-\lambda I)^{k+l+1} \implies v \in \ker(L-\lambda I)^{k+l}$$
- It follows that if
$$\ker((L-\lambda I)^k) = \ker((L-\lambda I)^{k+1}),$$
then for all $j \geq 0$
$$\ker((L-\lambda I)^k) = \ker((L-\lambda I)^{k+j})$$